

University of Groningen

Screening for Depression in Daily Life

Minaeva, Olga; Riese, Harriëtte; Lamers, Femke; Antypa, Niki; Wichers, Marieke; Booij, Sanne H

Published in:
Journal of Medical Internet Research

DOI:
[10.2196/22634](https://doi.org/10.2196/22634)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Minaeva, O., Riese, H., Lamers, F., Antypa, N., Wichers, M., & Booij, S. H. (2020). Screening for Depression in Daily Life: Development and External Validation of a Prediction Model Based on Actigraphy and Experience Sampling Method. *Journal of Medical Internet Research*, 22(12), [e22634].
<https://doi.org/10.2196/22634>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Original Paper

Screening for Depression in Daily Life: Development and External Validation of a Prediction Model Based on Actigraphy and Experience Sampling Method

Olga Minaeva¹, MSc; Harriëtte Riese¹, PhD; Femke Lamers², PhD; Niki Antypa³, PhD; Marieke Wichers¹, PhD; Sanne H Booij^{4,5}, PhD

¹Interdisciplinary Center Psychopathology and Emotion regulation (ICPE), Department of Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, Netherlands

²Department of Psychiatry, Amsterdam UMC, Amsterdam Public Health Research Institute, Vrije Universiteit, Amsterdam, Netherlands

³Department of Clinical Psychology, Institute of Psychology, Leiden University, Leiden, Netherlands

⁴Interdisciplinary Center for Psychopathology and Emotion regulation, Department of Developmental Psychology, Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, Netherlands

⁵Center for Integrative Psychiatry, Lentis, Groningen, Netherlands

Corresponding Author:

Olga Minaeva, MSc

Interdisciplinary Center Psychopathology and Emotion regulation (ICPE), Department of Psychiatry

University Medical Center Groningen

University of Groningen

Hanzeplein 1

Groningen, 9713 GZ

Netherlands

Phone: 31 50 361 2065

Email: o.minaeva@umcg.nl

Abstract

Background: In many countries, depressed individuals often first visit primary care settings for consultation, but a considerable number of clinically depressed patients remain unidentified. Introducing additional screening tools may facilitate the diagnostic process.

Objective: This study aimed to examine whether experience sampling method (ESM)-based measures of depressive affect and behaviors can discriminate depressed from nondepressed individuals. In addition, the added value of actigraphy-based measures was examined.

Methods: We used data from 2 samples to develop and validate prediction models. The development data set included 14 days of ESM and continuous actigraphy of currently depressed (n=43) and nondepressed individuals (n=82). The validation data set included 30 days of ESM and continuous actigraphy of currently depressed (n=27) and nondepressed individuals (n=27). Backward stepwise logistic regression analysis was applied to build the prediction models. Performance of the models was assessed with goodness-of-fit indices, calibration curves, and discriminative ability (area under the receiver operating characteristic curve [AUC]).

Results: In the development data set, the discriminative ability was good for the actigraphy model (AUC=0.790) and excellent for both the ESM (AUC=0.991) and the combined-domains model (AUC=0.993). In the validation data set, the discriminative ability was reasonable for the actigraphy model (AUC=0.648) and excellent for both the ESM (AUC=0.891) and the combined-domains model (AUC=0.892).

Conclusions: ESM is a good diagnostic predictor and is easy to calculate, and it therefore holds promise for implementation in clinical practice. Actigraphy shows no added value to ESM as a diagnostic predictor but might still be useful when ESM use is restricted.

(*J Med Internet Res* 2020;22(12):e22634) doi: [10.2196/22634](https://doi.org/10.2196/22634)

KEYWORDS

actigraphy; activity tracker; depression; experience sampling method; prediction model; screening

Introduction

Depressive disorders represent a major public health concern as they are the most prevalent psychiatric disorders and a leading cause of disability worldwide [1,2]. In many countries, depressed individuals often first visit primary care settings when they seek help [3]. Even though most nondepressed individuals can be accurately excluded in primary care [4], a considerable number of clinically depressed patients remains unidentified [5]. Thus, general practitioners can correctly identify between 41.7% and 53.0% of cases of depression with a sensitivity between 41.3% and 59.0% and a specificity between 74.5% and 87.3% [4]. An additional challenge in the detection of depression arises because patients often present with undefined or somatic illness [6], resulting in depression going undetected and often untreated for a longer time period [7].

According to a meta-analysis on the clinical diagnosis of depression in primary care, the accuracy of identification of depression can be improved by prospective examination over an extended period [4]. Therefore, introducing additional screening tools that allow continuous monitoring during daily life may facilitate the diagnostic process and improve referral of depressed individuals to the right care providers. A good candidate for a screening tool that holds particular value for studying mood disorders is the experience sampling method (ESM) [8]. Most commonly delivered via a smartphone, ESM involves repeated, intensive sampling of respondents' current affect, experiences, and behaviors while they are engaged in their daily activities, in their natural environments [9]. Hypothetically, this might be an optimal way of detecting depression risk, as a person can repeatedly assess his/her affect and behaviors in daily life with minimal retrospective recall bias [8]. Previous studies have shown that higher levels of negative affect, which is commonly assessed with ESM, are strongly associated with depression [10]. Furthermore, this method comes closest to the advised method to do longitudinal assessments of depressive symptoms [11,12]. A problem with ESM, however, may be that it places too high a burden on the patient, leading to reduced compliance [9]. Hence, ESM as a screening tool may not be suitable for everyone, warranting the exploration of more passive ways of collecting data as well.

A potential candidate for depression screening that involves passive data collection is ambulatory assessment of actigraphy data from sensors, such as activity trackers. Such activity trackers are now widely used, and they provide ecologically valid data about behavior [13-16]. The data derived from activity trackers include patterns of sleep, physical activity, and circadian rest-activity rhythm (RAR). Alterations in these patterns have been found in depression [17-22] and contributed to objective differentiation of depression subtypes [23]. Further, these behavioral parameters are easily and passively measurable by actigraphy and do not require any invasive procedure or active participation from individuals. Therefore, they do not create an additional burden for an individual [15]. However, it is important to note that the alterations in these actigraphy patterns are not

specific for depression, since altered sleep, physical activity, and RAR are present in many other health conditions [24,25]. Therefore, most likely, actigraphy may only be used in addition to other measures when screening for depression. However, the predictive value of actigraphy-based measures, alone and in combination with other measures, for depression remains to be examined.

Sleep, physical activity, and RAR domains are associated with each other [26,27]. However, previous researchers who found associations between actigraphy data and depression included measures from only 1 or 2 of these domains (ie, studying physical activity only, sleep only, or circadian RAR and sleep) [20,21,28-30], with a rare exception [31]. Therefore, probably not all actigraphy-based measures that have been previously associated with depression will have a unique predictive value as part of a multidimensional screening tool. Thus, it is currently still unclear which combination of actigraphy measures are most strongly associated with depression risk.

Using both ESM and actigraphy approaches together for continuous and everyday monitoring of behavioral and affective aspects in depression could be a promising screening tool. While it is assumed that self-reports of depression-related affect and behaviors as assessed with ESM predict depression better than behavioral sensor data, this assumption has never been tested before. To our knowledge, there is only 1 recent study that attempted to predict depression by using both ESM and actigraphy data; however, it was focused on the elderly, had a smaller sample size (N=47), and had no external validation [32]. Currently, it is not yet clear how these approaches perform and if they can be used for screening purposes, both separately and in combination.

In this study, we examined (1) whether ESM-assessed depression-related affect and behavior could discriminate between depressed and nondepressed individuals, (2) whether actigraphy data could discriminate between depressed and nondepressed individuals, and (3) whether actigraphy has added value with respect to the use of ESM. Therefore, we compared the performance of the prediction models with ESM only and actigraphy only to assess the added benefit of the individual domains and then evaluated the performance of the prediction model with both domains included. First, we hypothesized that ESM measures would have a better discriminating ability in distinguishing individuals with and without a diagnosis of depression than actigraphy measures. Second, adding actigraphy measures to the ESM prediction model would improve the discriminating ability. To test these hypotheses, we used 2 data sets for development and validation of the prediction models.

Methods**Study Population**

We used data from the Netherlands Study of Depression and Anxiety (NESDA) [33] to develop prediction models, and data

from the Mood and Movement in Daily Life (MOOVD) study to validate them [34].

Development Data Set

In short, NESDA is an ongoing multisite longitudinal cohort study among 2981 adults (aged 18 to 65 years) at baseline, including individuals with depressive and/or anxiety disorders and healthy control subjects, which were recruited from the general population. Details about the total NESDA sample are provided elsewhere [33]. In this study, we used a subsample from the Ecological Momentary Assessment and Actigraphy (EMAA) substudy, which combined 14 days of ESM (5 times a day) with continuous actigraphy [31,35]. A flowchart of the inclusion process for this study is provided in [Multimedia Appendix 1](#). Individuals with a diagnosed episode of major depressive disorder and/or dysthymia in the past month ($n=43$) and individuals with no lifetime depressive or anxiety disorder (control group, $n=82$) based on the Composite International Diagnostic Interview (CIDI) [36] were included in the study. Severity of depressive symptoms was assessed with the self-reported Inventory of Depressive Symptomatology (IDS-SR) [37]; the mean IDS-SR score represents moderate depressive symptoms.

Validation Data Set

The MOOVD study is an ambulatory assessment study among matched depressed and nondepressed individuals ($n=54$; aged 20 to 50 years) [34]. Depressed individuals were recruited from 3 psychiatric outpatient centers; nondepressed individuals were recruited from the general population in the Netherlands. This study combined 30 days of ESM (3 times a day) with continuous actigraphy. Depressed individuals ($n=27$) with a major depressive episode at the time of the interview or within two months prior to the interview, according to the CIDI, were included. Nondepressed individuals ($n=27$) were free of any mood disorders at the moment of inclusion but were allowed to have a history of depression ($n=1$, >7 years ago). Severity of depressive symptoms was assessed with the Beck Depression Inventory-II (BDI-II) [38]; the mean BDI-II score represents severe depressive symptoms. Individual scores, however, range from no/mild to severe depressive symptoms in both data sets (IDS-SR score between 9 and 64; BDI-II score between 15 and 51).

Actigraphy Assessments

NESDA participants wore the wrist-worn GENEActiv accelerometer (Activinsights Ltd) for 24 hours a day for 14 days. GENEActiv validity studies have demonstrated strong correlations for criterion validity (Pearson $r=0.79-0.98$) [39] and a good ability to determine sedentary behavior in adults (aged 18 to 55 years) (Pearson $r=0.81$) [40]. Details of the actigraphy measurements of the NESDA-EMAA substudy are provided elsewhere [31]. MOOVD participants were assessed with an Actical accelerometer (Respironics, Inc) for 24 hours a day for 30 days. In the laboratory study, the Actical demonstrated high reliability (intraclass correlation coefficient=0.92) and validity ($r=0.81$) in adolescents [41]. More information about the actigraphy assessments of the MOOVD study can be found elsewhere [42].

Experience Sampling Methodology (ESM)

In NESDA, participants took part in the ESM assessment for 2 weeks, during which they filled out questions on smartphones 5 times a day. The electronic diary had a fixed design with 3 hours between each beep, and the questionnaire included items on current mood states, social interactions, daily experiences, and behaviors [35]. Of all ESM assessments of all participants, only 8.3% were missing and all included participants had enough valid data points (>60 time points). In the MOOVD study, participants completed questionnaires on an electronic diary, the PsyMate (PsyMate BV) [43]. The electronic diary had a fixed design with 3 beeps a day, 6 hours apart. The electronic questionnaire contained items about mood, sleep, activities, as well as social interactions, important events, rumination, and self-esteem. Detailed information about the ambulatory assessment procedure is provided elsewhere [34]. All included participants had enough valid diary measurements (>60 time points).

Outcome Variables

As the main outcome measure for both data sets, we used presence or absence of a diagnosis of depression (major depressive disorder and/or dysthymia) based on DSM-IV criteria [44], assessed with the Composite International Diagnostic Interview (CIDI), version 2.1. The CIDI is a fully structured interview designed for assessing mental disorders according to the diagnostic criteria of the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) [36]. All CIDs were performed by well-trained research assistants, mainly psychologists, mental health care nurses, and residents in psychiatry. In the development data set, participants were diagnosed with the CIDI instrument during the regular NESDA interview wave, which was a maximum of 31 days prior to the actigraphy and ESM assessments. In the validation data set, participants started the actigraphy and ESM assessments immediately following the screening CIDI.

Predictor Variables

Objectively assessed sleep, physical activity, and computed RAR variables, as calculated from the actigraphy data, and ESM-assessed depression-related affect and behavior were used as predictors in our models. Preprocessing of the raw actigraphy data was done in R using GGIR package version 1.5-18 (<https://cran.r-project.org/web/packages/GGIR/>) for the NESDA data set and described elsewhere [31]. Almost all variables were created similarly in the NESDA and the MOOVD data sets; any exceptions are mentioned.

In the NESDA data set, physical activity was assessed as gross motor activity per day and as minutes of moderate-to-vigorous physical activity (MVPA) per day [45]. Objective gross motor activity was estimated by calculating the Euclidian Norm Minus One (ENMO) per individual per day [31]. Based on those calculations, average estimates of gross motor activity were estimated for each participant. To keep consistent with earlier papers, MVPA was defined as ENMO values greater than 125 mg [31]. In the MOOVD data set, the Actical actigraphy device did not allow the extraction of the raw actigraphy data (data in SI units represented as acceleration in x, y, and z axes), and

therefore the data were not processed with the GGIR package. Instead, activity counts (AC) and activity intensity with 4 categories (sedentary, light, moderate, and vigorous) were calculated as a measure of motor activity by an in-built algorithm of the Actical software.

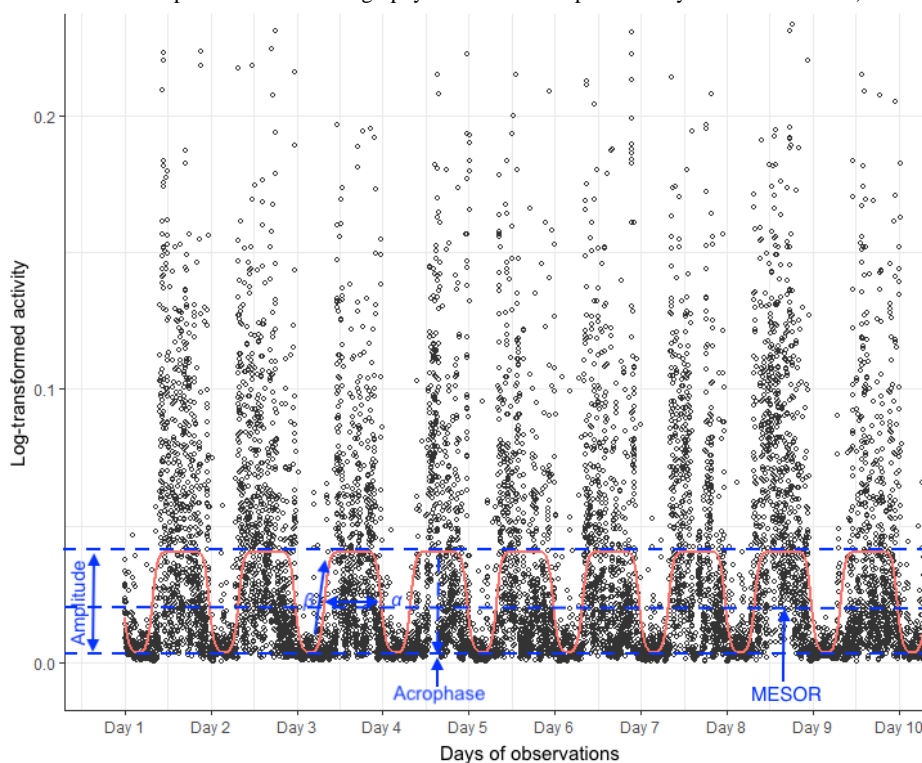
Sleep was assessed as total sleep time (TST) in hours and sleep efficiency per night (%) [21,46]. In the NESDA data set, TST was estimated with the GGIR package and equaled the accumulated nocturnal sustained inactivity bouts. Sleep efficiency was calculated as TST divided by time in bed (estimated by the GGIR package). For the MOOVD data set, TST was calculated as the sum of estimated sleep periods based on the Sadeh algorithm [47]. Sleep efficiency was calculated as a percentage of time scored as sleep during the time spent in bed.

To characterize circadian RAR profiles, individual actigraphy data sets were fitted to an extended cosinor model [48] using nonlinear least-squares regression (RAR package version 1.0.0 for R). This allowed the estimation of 5 circadian curve parameters for each participant, namely the midline estimating

statistic of rhythm (MESOR), amplitude, acrophase, α , and β , as well as the circadian rhythmicity index (F statistic) (see Figure 1 for more details).

To assess depression-related affect and behavior, we selected ESM items that, in terms of content, matched DSM-V diagnostic criteria for depression. For example, the symptom of “sad or depressed mood” could be represented by the momentary affect state, “I feel sad.” The following items that were present in both data sets were included: sad or depressed mood, irritation, appetite, energy, tiredness, loss of interest, enthusiasm, guilt, concentration, and sleep disturbances. A complete list of included ESM items from the NESDA and the MOOVD data sets can be found in Multimedia Appendix 2. All items were scored on 7-point Likert scales. The sum score of these items calculated for each day and then averaged across 14 days represented depression-related affect and behavior. The person-level Cronbach α for depression-related affect and behavior was .928 in the NESDA data set and .936 in the MOOVD data set. Gender, age, and education level were included in the analysis as covariates.

Figure 1. Example of rest-activity rhythm parameters derived from the extended cosinor model. The midline estimating statistic of rhythm (MESOR) is a mean of the modeled activity curve; amplitude is the difference between the peak and trough of the fitted curve, herein estimating the range of activity levels across the 24-hour period; acrophase is a phase marker indicating the time when the fitted curve reaches its peak (ie, time of maximal activity levels across the 24-hour period); α is the relative width of the curve at the middle of the peak; β is an indicator of the steepness of the rise and fall of the curve; and circadian rhythmicity index (F statistic) is an indicator of the strength of circadian rhythmicity (a goodness-of-fit measure for which higher values indicate smaller discrepancies between actigraphy data and values predicted by the cosinor model).



Statistical Analysis

Multicollinearity for all predictor variables was checked by calculating Spearman correlations and the variance inflation factors (VIFs). Spearman correlations above 0.80 and VIFs above 10 were considered to be indicative of severe collinearity [49,50]. In this situation, 1 of the collinear variables that was the least related to the outcome variable was removed from

further analysis. Fractional polynomials were used to check the presence of nonlinear associations of the continuous predictors to the outcome variable. Cubic association was found for sleep duration (TST) and therefore was included as such in the analysis.

Building Single-Domain Models

The next step was to build single-domain models for actigraphy and ESM measurements separately. For the ESM model (1), we included the ESM “depression-related affect and behavior” score and covariates (age, gender, and education), as their association to depression has been consistently shown [51,52] and this information can be easily added to a screening tool. The sum score was chosen instead of including the ESM items in the analysis separately, as it was meant to mimic depressive symptoms.

$$\text{Group status} = a_0 + a_1 \text{sum score} + a_2 \text{age} + a_3 \text{gender} + a_4 \text{education} \quad (1)$$

where “a” represents the regression coefficients from the model: a_1 - a_4 are predictor coefficients and a_0 is the intercept.

To build an actigraphy model, we used a multivariate backward stepwise logistic regression approach [53]. A baseline actigraphy model included all actigraphy predictors (amplitude, acrophase, α , β , F statistic, ENMO, TST, and sleep efficiency) as well as predefined covariates (age, gender, and education). The MESOR was found to be collinear with ENMO and the least related to the outcome variable; therefore, it was removed from the further analysis. Since different physical activity metrics (ENMO and AC) were available for the development and the validation data sets, we standardized ENMO and AC to alleviate a comparison of the actigraphy models in two data sets. In the following steps, we removed the least significant actigraphy variable (with the highest P value) and compared the Akaike information criterion (AIC) value to the AIC from the previous model. A significantly smaller AIC indicates a better model. The procedure was repeated until we defined the optimal combination of actigraphy predictors based on the AIC. The regression equation for the final actigraphy model is included below (2):

$$\text{Group status} = a_0 + a_1 \text{standardized ENMO} + a_2 \text{acrophase} + a_3 \text{age} + a_4 \text{gender} + a_5 \text{education} \quad (2)$$

Building a Combined-Domains Model

A multivariable backward stepwise logistic regression model was performed to examine what combination of predictors (ie, actigraphy, ESM) resulted in the optimal prediction model for distinguishing between depressed and nondepressed individuals. Since we used a backward approach, the baseline model included all predictors with unique information based on the single-domain models (the actigraphy and the ESM models). In the following models, we used a procedure where we removed

variables one by one, based on the highest P value, and checked every time whether AIC improved until we defined the final prediction model based on the AIC [54]. The regression equation for the final combined-domains model is included below (3):

$$\text{Group status} = a_0 + a_1 \text{sum score} + a_2 \text{standardized ENMO} + a_3 \text{age} + a_4 \text{gender} + a_5 \text{education} \quad (3)$$

Evaluation and Validation of the Single-Domain and the Combined-Domains Prediction Models

To evaluate the performance of the combined-domains model and the single-domain models, we utilized goodness-of-fit indices and calibration curves and assessed the discriminative ability of the models (the area under the receiver operating characteristic curve [AUC]) [55]. The goodness-of-fit indices and calibration curves evaluate how close the predicted and observed estimates are. The AUC represents the ability of the models to distinguish between patients with and without the depression diagnosis and ranges from 0.5 (by chance) to 1.0 (perfect discrimination). These quality indicators of all 3 models were compared with a basic model with only the covariates and to each other. As suggested in the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) statement [56], we performed bootstrapping techniques for internal validation of the model to simulate the performance of the prediction model in comparable patient data sets.

As a next step, we performed an external validation by using the developed single-domain and combined-domains models from the NESDA data set to assess the predictive performance of the models in the validation sample (ie, the MOOVD data set), calculating the discriminative ability reflected by the AUC. The NESDA data set was chosen to be a development data set because it was larger than the MOOVD data set, to minimize the possibility of overfitting while building the prediction model. The MOOVD data set was suitable for the external validation because inclusion criteria and measurements of depression were similar. Again, we compared the final combined-domains prediction model with the actigraphy and ESM models in the MOOVD data set to check whether it still had a better fit than the single-domain models. The results were reported according to the TRIPOD statement [56].

Results

The characteristics of the development (NESDA) and validation (MOOVD) data sets are given in Table 1.

Table 1. Characteristics of development and validation data sets.

	NESDA ^a (n=125)		MOOVD ^b (n=54)	
	Depressed (n=43)	Control (n=82)	Depressed (n=27)	Control (n=27)
Study characteristics				
Data collection period	2014-2017		2012-2014	
Setting	General population, primary health care, and mental health care		General population, and outpatient centers for psychiatry	
Inclusion criteria for cases	Receiving a depression diagnosis 1 month before the ESM ^c /actigraphy assessment		Receiving a depression diagnosis 1 month before the ESM/actigraphy assessment	
Outcome	Presence or absence of a depression diagnosis (MDD ^d and/or dysthymia) based on DSM-IV ^e criteria		Presence or absence of a depression diagnosis (MDD and/or dysthymia) based on DSM-IV criteria	
Prevalence of outcome, n (%)	43 (34.4)		27 (50.0)	
Sociodemographic characteristics				
Female, n (%)	29 (67.4)	46 (56.1)	20 (74.1)	20 (74.1)
Age, mean (SD)	52.14 (9.57)	51.50 (12.70)	34.70 (9.86)	34.04 (8.96)
Education (high), n (%)	13 (30.2)	49 (59.8)	17 (63.0)	19 (70.4)
Psychopathology				
Depression severity instrument	IDS-SR ^f		BDI-II ^g	
Depression severity, mean (range)	34.53 (9-64)	5.44 (0-25)	31.33 (15-51)	2.26 (0-10)
AD ^h and/or BD ⁱ use, n (%)	23 (53.5)	4 (4.9)	15 (55.6)	1 (3.7)

^aNESDA: the Netherlands Study of Depression and Anxiety.^bMOOVD: Mood and Movement in Daily Life.^cESM: experience sampling method.^dMDD: major depressive disorder.^eDSM-IV: Diagnostic and Statistical Manual of Mental Disorders, fourth edition.^fIDS-SR: Inventory of Depressive Symptomatology (self-report).^gBDI-II: Beck Depression Inventory-II.^hAD: antidepressant.ⁱBD: benzodiazepine.

Predictors of depression in the final models were the “depression-related affect and behavior” sum score for the ESM model and gross motor activity (ENMO) and the time of maximal activity levels across the 24-hour period (acrophase) for the actigraphy model (Table 2). The combined-domains prediction model included the “depression-related affect and behavior” sum score and ENMO variables.

For the ESM model, the predictive capacity was 95.2%, which is 29.6% higher than in the null (only intercept) model (65.6%). Calibration of the ESM model was adequate with a Nagelkerke R^2 statistic of 0.904 and a Hosmer-Lemeshow goodness-of-fit test of 0.960 ($P=.998$).

For the actigraphy model, the predictive capacity of the final step of the backward selection model was 71.8% (6.2% higher than the null model). Calibration of the actigraphy model was adequate with a Nagelkerke R^2 statistic of 0.357 and a Hosmer-Lemeshow goodness-of-fit test of 10.678 ($P=.221$).

The final (combined-domains) model had the same predictive capacity as the ESM model (95.2%). Calibration of the final model was adequate with a Nagelkerke R^2 statistic of 0.913 and a Hosmer-Lemeshow goodness-of-fit test of 1.786 ($P=.987$). In all 3 calibration plots, the slope approached the diagonal (see Multimedia Appendix 3).

Table 2. Predictors of depression included in the experience sampling method (ESM) prediction model, the actigraphy prediction model, and the final combined-domains model.

Predictor	B ^a	SE	P value	Exp(B) ^b	95% CI for Exp(B)
ESM prediction model					
Intercept	-27.880	8.668	.001 ^c	0.000	
Gender	-2.955	1.597	.06	0.052	0.002-1.191
Age	0.143	0.068	.03 ^c	1.154	1.009-1.319
Education	-1.701	1.058	.11	0.182	0.023-1.452
Sum score ^d	0.891	0.243	<.001 ^c	2.437	1.513-3.923
Actigraphy prediction model					
Intercept	0.050	2.954	.99	1.052	
Gender	-1.021	0.513	.046 ^c	0.360	0.132-0.984
Age	-0.012	0.020	.54	0.988	0.950-1.027
Education	-1.711	0.456	.001 ^c	0.181	0.074-0.441
Z-score ENMO ^e	-1.004	0.283	<.001 ^c	0.367	0.211-0.638
Acrophase	0.289	0.177	.10	1.335	0.943-1.888
Final combined-domains prediction model					
Intercept	-25.841	9.314	.01 ^c	0.000	
Gender	-3.751	1.740	.03 ^c	0.024	0.001-0.712
Age	0.102	0.076	.18	1.107	0.953-1.286
Education	-1.988	1.163	.09	0.137	0.014-1.339
Sum score	0.919	0.263	<.001 ^c	2.508	1.498-4.197
Z-score ENMO	-1.117	0.767	.15	0.327	0.073-1.472

^aB: regression coefficient.^bExp(B): exponentiation of the B coefficient (odds ratio).^cSignificant P value ($P < .05$).^dSum score: a person-level sum score of the ESM items that represent depression-related affect and behavior.^eENMO: Euclidian Norm Minus One.

In the development data set, the discriminative ability in predicting depression was good in the actigraphy model (AUC=0.790) and excellent in the ESM model (AUC=0.991) and the combined-domains model (AUC=0.993) (Table 3, Figure 2). The estimations did not meaningfully change after internal validation. The discriminative ability of the model in the external validation sample was reasonable in the actigraphy model (AUC=0.648) and very good in the ESM model (AUC=0.891)

and the combined-domains model (AUC=0.892) (Multimedia Appendix 4). Calibration of all models was adequate, with the slope approaching the diagonal in all 3 calibration plots (Multimedia Appendix 5). Multimedia Appendices 6 and 7 provide an overview of different probability thresholds and their respective classification measures (sensitivity, specificity, and predictive values) for all 3 models of interest in the development and validation data sets.

Table 3. The discriminative ability of the final combined-domains model and the basic, actigraphy, and experience sampling method (ESM) models.

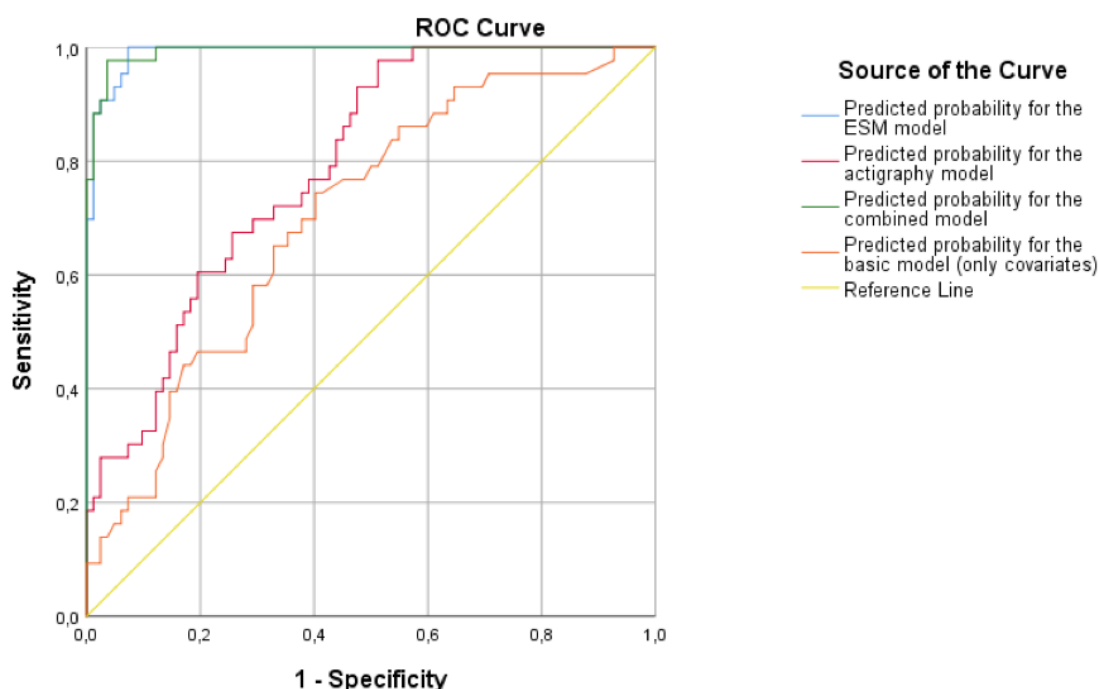
Models	Development data set (NESDA ^a ; n=125)		Validation data set (MOOVD ^b ; n=54)	
	AUC ^c	95% CI	AUC	95% CI
Basic model ^d	0.704	0.610-0.798	0.464	0.304-0.624
ESM model	0.991	0.981-1.000	0.891	0.802-0.979
Actigraphy model	0.790	0.713-0.867	0.648	0.492-0.803
Combined-domains model	0.993	0.983-1.000	0.892	0.800-0.984

^aNESDA: the Netherlands Study of Depression and Anxiety.

^bMOOVD: Mood and Movement in Daily Life.

^cAUC: area under the receiver operating characteristic curve.

^dBasic model has covariates included only (gender, age, and education).

Figure 2. The receiver operating characteristic (ROC) curves of the basic model, the experience sampling method (ESM) model, the actigraphy model, and the combined-domains model in the development data set (the Netherlands Study of Depression and Anxiety).

Discussion

Principal Results

In this paper, we have developed 3 prediction models based on ESM measures of depression-related affect and behaviors, on actigraphy, and on their combination, for discriminating currently depressed from nondepressed individuals. To our knowledge, this is the first study that has created and compared such models for their individual performance and for their performance in combination, and using both internal and external validation to test their discriminative abilities. The ESM model had an excellent predictive potential in discriminating depressed and nondepressed individuals in both the development and the validation data sets. The actigraphy model, in turn, had a reasonable predictive potential alone but could not compete with the ESM model in predictive performance. The combined-domains prediction model, which

included the ESM measure as well as the best combination of actigraphy measures, was very similar to the ESM model in performance in both the development and the validation data sets. Hence, from the results we can conclude that the ESM and actigraphy measures both have the potential to serve as an additional screening tool; however, actigraphy does not have added value when combined with ESM.

Comparison with Prior Work

The ESM sum score combined items about positive affect, negative affect, sleep, and appetite. Measuring these symptom-related items over a prolonged period of time resulted in successful discrimination between depressed and nondepressed individuals. The constructed prediction model performed excellently, not only in the development but also in the validation data set. In line with our findings, multiple previous studies showed correlations between related measures, mainly negative affect and positive affect (assessed with ESM),

and depression [10,12,57]. Another study attempted to estimate depressive symptoms based on the ESM items and found significant correlations with depressive symptoms assessed with symptom questionnaires [11]. Despite the strong correlation between ESM-assessed depression-related affect and behavior and depression, there have been no diagnostic prediction studies using such ESM data. Existing literature, however, indicates that both negative affect and positive affect play a role in predicting relapse of depression [58,59] and future treatment outcome [60,61]. Additionally, negative affect was found to be predictive of depression onset in youth [62]. There are a very limited number of prognostic studies available, focusing on older adults [63]. Authors have identified the items “sad” and “tired” as sensitive measures that have the potential to predict future depression status in older adults. Although replication in other types of samples is warranted, our results suggest that the ESM measure has potential for screening purposes in clinical practice.

Interestingly, self-report measures of depressive affect and behavior largely overperformed objective measures of behavior in distinguishing depressed from nondepressed individuals. Researchers often make the implicit assumption that subjective measures are inaccurate and cannot compete with objective assessments. Self-reported affect and behavior, indeed, might not align with actual observed affect or behavior due to a difference between perceived and actual behavior [64]. Nevertheless, this bias in how a depressed person perceives their own emotions and feelings might be highly useful for predicting depression.

Although the actigraphy model showed good performance in the development data set, its performance dropped in the validation data set. While this discrepancy may be a signal of worse performance in other samples, it could also partly be due to differences in physical activity metrics in use. We found that metrics from the GENEActiv and Actical accelerometers (ENMO and AC, respectively) have never been directly compared before, hence we could not apply any known formula to transform data from one into the other. To overcome this issue, we z-transformed ENMO and AC before the analysis to adjust the variables to the same scale from -1 to 1 . This step improved the performance of the actigraphy model, although it was still lower in the validation data set. Of note, it has previously been shown that the placement of the accelerometer device on the body influences its outcome measures in both the level of and fluctuations in activities [65]. Wrist placement was recommended as a basic technique to capture motor activity in depressed patients because it records whole-body movement and gestures [66]. In this study, the devices were worn on the wrist, which implies that the model is only applicable to actigraphy as measured using the wrist placement. Despite the existing challenges and a lower predictive capacity compared with ESM, actigraphy, being a passive data collection method, might still be useful for screening purposes in some cases, for example, in a situation where the use of ESM is not possible or too bothersome for particular individuals.

In agreement with previous studies that used actigraphy data [20-22,31,67], we have found strong associations between lower levels of physical activity and depressive symptoms/depressive

disorders. Other RAR characteristics such as MESOR, amplitude, α , β , and F statistic, calculated with the extended cosinor analysis, were not sufficiently predictive when assessed together with the overall activity level (ENMO) and therefore, were eliminated from the model. Interestingly, the individual associations between these parameters and depression were significant; however, when these variables were combined, only a few remained in the model (ENMO and acrophase). The fact that only a few variables remained might be due to the potential overlap of these parameters. Even though the association between circadian rest-activity parameters and depression has been shown previously [30,68-71], to our knowledge, there are no studies that specifically compared the predictive ability of various actigraphy-based parameters from different domains in distinguishing depressed and non-depressed individuals. Hence, future studies should further examine to what extent these different variables are measuring partly the same concepts.

The combined-domains model had an excellent performance that was highly similar to the performance of the ESM model. Adding the actigraphy model component (ENMO or AC) did not significantly improve the performance of the combined-domains model. The explanation for this lack of improvement might be that ESM captures part of actigraphy variance. This overlap may be in part because both ESM and actigraphy assess sleep; one assesses subjective sleep duration and subjective sleep quality, while the other assesses the same characteristics objectively. Another example of shared variance might be that complaints like anhedonia could result in a patient being less active and hence, the reduced physical activity may be a result of the depressive symptoms. As in the previous example, concentration problems could be associated with lower sleep quality [72]. This makes questionable the hypothesis of whether actigraphy has added value in its own right when combined with ESM. To our knowledge, there is only 1 recent study that attempted to predict depression by using both ESM and actigraphy data; however, it was focused on community-dwelling older adults, had a smaller sample size ($N=47$), and had no external validation [32]. These researchers used a wide range of various predictors based on actigraphy and ESM in a machine learning approach to define the optimal combination of the predictors and build a prediction model. As in our study, these researchers developed their model on the basis of daily mean ESM scores, and actigraphy-based daily mean activity levels and daily sleep efficiency variables, although the latter was removed from our prediction models. Hence, the chosen variables showed predictive potential in detecting depression even when different selection approaches were applied. More details of the studies discussed in this paragraph are provided in [Multimedia Appendix 8](#).

Strengths and Limitations

The main strength of this study was the external validation of the prediction models in a different and adequate data set that prevented overstating the results [56,73]. The fact that there were substantial differences between the 2 data sets and the results were consistent in both data sets provides strong evidence that the prediction models can be generalized to new patients [73]. Finally, all constructed models include variables that are ecologically valid and easy to measure in real life. There are

also some limitations of the study that need to be mentioned. First, different physical activity metrics (ENMO versus AC) limited the ability to externally validate the actigraphy prediction model in an optimal way. To avoid this problem in the future, researchers should preferably use the accelerometers that allow access to the raw data. In this case, the same output metric can be chosen so that the algorithms for the computation will be the same or at least comparable. Second, the development data set had a small number of individuals in the depression groups who reported no or mild depressive symptoms. This was most likely due to logistic reasons typical of large cohort studies, as our study was. Participants were diagnosed with the CIDI instrument during the regular NESDA interview wave, which was a maximum of 31 days prior to the actigraphy assessments, whereas depression severity assessment with the IDS-SR questionnaire was not necessarily done close to the actigraphy assessment period (up to 72 days prior and two cases of 74 and 351 days after the CIDI). This situation, however, could possibly deflate the association rather than inflate it, and the model would have more predictive capacity in more acutely depressed individuals. Third, the size of the validation data set was smaller than suggested by some simulation studies, which required at least 100 events for the validation sample. This suggestion, however, was based on limited simulation studies and it lacks the empirical evidence to guide research [56]. The sample size, therefore, is often determined by the available data, as was the case in our study [56]. Finally, both development and validation samples were collected in the Netherlands, which limits generalizability to countries with similar ethnicity profiles and health care systems.

Concerning future research, the constructed ESM models performed excellently with 14 and 30 days of continuous measurements in the development and validation data sets, respectively. However, if a shorter duration of ESM measurements showed similar performance in the prediction of depression, a shorter measurement period could potentially reduce the burden on patients who struggle with longer ESM regimes. Indeed, it has previously been shown that an association between negative affect and depression can already be detected using ESM for 6 to 7 days [60,74,75] or even 3 days [76]. Further, different accelerometer devices can differ substantially in sampling frequency, data processing algorithms, and other characteristics [77]. Therefore, examining various accelerometers and possibly creating formulas for output

transformation might be valuable to facilitate the comparison of outputs from different devices. In this study, we assessed 2 samples with currently depressed and nondepressed individuals with moderate to severe levels of depressive symptoms. Therefore, it remains to be seen whether the results of the constructed models generalize to a population with more ambiguous depressive symptoms or other psychiatric problems. A large-scale study in the general population is needed before recommendations can be made for the use of such prediction models in clinical practice. If this is the case, implementation of such a prediction tool in practice can be relatively straightforward. First, ESM seems to be a feasible tool for clinical practice and has the additional benefit that clients become co-owners of the care process [78,79]. Second, the developed algorithm for calculating “depression-related affect and behavior” sum scores is easy to use. However, further automation processes are needed to facilitate use of such a tool in real-life settings, such as primary care. Finally, such studies, including the current one, are necessary to build up empirical evidence on mechanisms, which will be the basis for developing medical devices in the future. Using strict regulations for such kinds of basic research might strangle innovative studies rather than pushing them forward.

Conclusions

To conclude, in this study we presented models to predict depression based on ESM-assessed depression-related affect and behavior as well as actigraphy data. The most potent predictor was the “depression-related affect and behavior” sum score constructed from the ESM data. The ESM model had an excellent predictive capacity, is easy to calculate, and hence might be feasible for future implementation in clinical practice. The actigraphy model had reasonable performance, but there was no added value of actigraphy in combination with ESM. Despite the fact that the actigraphy model had a lower predictive capacity than ESM, actigraphy could still be of value in situations where ESM is too burdensome.

Authors' Contributors

OM, SHB, MW, and HR were involved in the design of the study. OM and SHB performed the statistical analysis. OM wrote the original draft of the manuscript. OM, SHB, MW, HR, FL, and NA critically revised the draft for intellectual content and approved the final version.

Acknowledgments

The infrastructure for NESDA is funded through the Geestkracht program of the Netherlands Organization for Health Research and Development (ZonMw, grant number 10-000-1002) and financial contributions by participating universities and mental health care organizations (VU University Medical Center, GGZ inGeest, Leiden University Medical Center, Leiden University, GGZ Rivierduinen, University Medical Center Groningen, University of Groningen, Lentis, GGZ Friesland, GGZ Drenthe, Rob Giel Onderzoekscentrum). The authors thank all NESDA and MOOVD participants, research assistants, and students who made the data acquisition possible.

This study was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovative programme (ERC-CoG-2015; No 681466 to M Wichers).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Flowchart of a study sample from the Netherland Study of Depression and Anxiety-Ecological Momentary Assessment and Actigraphy (NESDA-EMAA) study.

[\[DOCX File , 32 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

List of included experience sampling method (ESM) items from the Netherland Study of Depression and Anxiety (NESDA) and the Mood and Movement in Daily Life (MOOVD) data sets and corresponding Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) criteria.

[\[DOCX File , 17 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Calibration plots of the models in the development data set.

[\[DOCX File , 86 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

The receiver operating characteristic (ROC) curves in the validation dataset.

[\[DOCX File , 35 KB-Multimedia Appendix 4\]](#)

Multimedia Appendix 5

Calibration plots of the models in the validation data set.

[\[DOCX File , 80 KB-Multimedia Appendix 5\]](#)

Multimedia Appendix 6

Sensitivity, specificity, and cutoff score for the experience sampling method (ESM) model, the actigraphy model, and the final (combined-domains) model in the development data set.

[\[DOCX File , 16 KB-Multimedia Appendix 6\]](#)

Multimedia Appendix 7

Sensitivity, specificity, and cutoff score for the experience sampling method (ESM) model, the actigraphy model, and the final (combined-domains) model in the validation data set.

[\[DOCX File , 13 KB-Multimedia Appendix 7\]](#)

Multimedia Appendix 8

Quantitative comparisons of the articles referred to in the discussion.

[\[DOCX File , 21 KB-Multimedia Appendix 8\]](#)

References

1. GBD 2017 DiseaseInjury IncidencePrevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018 Nov 10;392(10159):1789-1858 [[FREE Full text](#)] [doi: [10.1016/S0140-6736\(18\)32279-7](https://doi.org/10.1016/S0140-6736(18)32279-7)] [Medline: [30496104](https://pubmed.ncbi.nlm.nih.gov/30496104/)]
2. Depression. World Health Organization. 2019. URL: <https://www.who.int/news-room/fact-sheets/detail/depression> [accessed 2020-01-17]
3. van Weel C, van Weel-Baumgarten E, van Rijswijk E. Treatment of depression in primary care. *BMJ* 2009 Mar 19;338:b934. [doi: [10.1136/bmj.b934](https://doi.org/10.1136/bmj.b934)] [Medline: [19299476](https://pubmed.ncbi.nlm.nih.gov/19299476/)]
4. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet* 2009 Aug 22;374(9690):609-619. [doi: [10.1016/S0140-6736\(09\)60879-5](https://doi.org/10.1016/S0140-6736(09)60879-5)] [Medline: [19640579](https://pubmed.ncbi.nlm.nih.gov/19640579/)]
5. Fernández A, Pinto-Meza A, Bellón JA, Roura-Poch P, Haro JM, Autonell J, et al. Is major depression adequately diagnosed and treated by general practitioners? Results from an epidemiological study. *Gen Hosp Psychiatry* 2010;32(2):201-209. [doi: [10.1016/j.genhosppsych.2009.11.015](https://doi.org/10.1016/j.genhosppsych.2009.11.015)] [Medline: [20302995](https://pubmed.ncbi.nlm.nih.gov/20302995/)]
6. Carey M, Jones K, Meadows G, Sanson-Fisher R, D'Este C, Inder K, et al. Accuracy of general practitioner unassisted detection of depression. *Aust N Z J Psychiatry* 2014 Jan 10;48(6):571-578 [[FREE Full text](#)] [doi: [10.1177/0004867413520047](https://doi.org/10.1177/0004867413520047)] [Medline: [24413807](https://pubmed.ncbi.nlm.nih.gov/24413807/)]

7. Thornicroft G. No time to lose: onset and treatment delay for mental disorders. *Epidemiol Psychiatr Sci* 2012 Mar;21(1):59-61. [doi: [10.1017/s2045796011000825](https://doi.org/10.1017/s2045796011000825)] [Medline: [22670413](https://pubmed.ncbi.nlm.nih.gov/22670413/)]
8. Wenze SJ, Miller IW. Use of ecological momentary assessment in mood disorders research. *Clin Psychol Rev* 2010 Aug;30(6):794-804. [doi: [10.1016/j.cpr.2010.06.007](https://doi.org/10.1016/j.cpr.2010.06.007)] [Medline: [20619520](https://pubmed.ncbi.nlm.nih.gov/20619520/)]
9. Ebner-Priemer UW, Trull TJ. Ecological momentary assessment of mood disorders and mood dysregulation. *Psychol Assess* 2009 Dec;21(4):463-475. [doi: [10.1037/a0017075](https://doi.org/10.1037/a0017075)] [Medline: [19947781](https://pubmed.ncbi.nlm.nih.gov/19947781/)]
10. aan het Rot M, Hogenelst K, Schoevers RA. Mood disorders in everyday life: a systematic review of experience sampling and ecological momentary assessment studies. *Clin Psychol Rev* 2012 Aug;32(6):510-523. [doi: [10.1016/j.cpr.2012.05.007](https://doi.org/10.1016/j.cpr.2012.05.007)] [Medline: [22721999](https://pubmed.ncbi.nlm.nih.gov/22721999/)]
11. Arney MF, Schatten HT, Haradhvala N, Miller IW. Ecological momentary assessment (EMA) of depression-related phenomena. *Curr Opin Psychol* 2015 Aug 1;4:21-25. [doi: [10.1016/j.copsyc.2015.01.002](https://doi.org/10.1016/j.copsyc.2015.01.002)] [Medline: [25664334](https://pubmed.ncbi.nlm.nih.gov/25664334/)]
12. Telford C, McCarthy-Jones S, Corcoran R, Rowse G. Experience sampling methodology studies of depression: the state of the art. *Psychol Med* 2012 Jun;42(6):1119-1129. [doi: [10.1017/S0033291711002200](https://doi.org/10.1017/S0033291711002200)] [Medline: [22008511](https://pubmed.ncbi.nlm.nih.gov/22008511/)]
13. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. *Int J Behav Nutr Phys Act* 2015;12:42 [FREE Full text] [doi: [10.1186/s12966-015-0201-9](https://doi.org/10.1186/s12966-015-0201-9)] [Medline: [25890168](https://pubmed.ncbi.nlm.nih.gov/25890168/)]
14. van de Water ATM, Holmes A, Hurley DA. Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography - a systematic review. *J Sleep Res* 2011 Mar;20(1 Pt 2):183-200 [FREE Full text] [doi: [10.1111/j.1365-2869.2009.00814.x](https://doi.org/10.1111/j.1365-2869.2009.00814.x)] [Medline: [20374444](https://pubmed.ncbi.nlm.nih.gov/20374444/)]
15. Hills AP, Mokhtar N, Byrne NM. Assessment of physical activity and energy expenditure: an overview of objective measures. *Front Nutr* 2014;1:5 [FREE Full text] [doi: [10.3389/fnut.2014.00005](https://doi.org/10.3389/fnut.2014.00005)] [Medline: [25988109](https://pubmed.ncbi.nlm.nih.gov/25988109/)]
16. Kooiman TJM, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. *BMC Sports Sci Med Rehabil* 2015;7:24 [FREE Full text] [doi: [10.1186/s13102-015-0018-5](https://doi.org/10.1186/s13102-015-0018-5)] [Medline: [26464801](https://pubmed.ncbi.nlm.nih.gov/26464801/)]
17. Germain A, Kupfer DJ. Circadian rhythm disturbances in depression. *Hum Psychopharmacol* 2008 Oct;23(7):571-585 [FREE Full text] [doi: [10.1002/hup.964](https://doi.org/10.1002/hup.964)] [Medline: [18680211](https://pubmed.ncbi.nlm.nih.gov/18680211/)]
18. Hickie IB, Naismith SL, Robillard R, Scott EM, Hermens DF. Manipulating the sleep-wake cycle and circadian rhythms to improve clinical management of major depression. *BMC Med* 2013 Mar 22;11:79 [FREE Full text] [doi: [10.1186/1741-7015-11-79](https://doi.org/10.1186/1741-7015-11-79)] [Medline: [23521808](https://pubmed.ncbi.nlm.nih.gov/23521808/)]
19. Vadnie CA, McClung CA. Circadian rhythm disturbances in mood disorders: insights into the role of the suprachiasmatic nucleus. *Neural Plast* 2017;2017:1504507 [FREE Full text] [doi: [10.1155/2017/1504507](https://doi.org/10.1155/2017/1504507)] [Medline: [29230328](https://pubmed.ncbi.nlm.nih.gov/29230328/)]
20. Burton C, McKinstry B, Szentagotai TA, Serrano-Blanco A, Pagliari C, Wolters M. Activity monitoring in patients with depression: a systematic review. *J Affect Disord* 2013 Feb 15;145(1):21-28. [doi: [10.1016/j.jad.2012.07.001](https://doi.org/10.1016/j.jad.2012.07.001)] [Medline: [22868056](https://pubmed.ncbi.nlm.nih.gov/22868056/)]
21. Hori H, Koga N, Hidese S, Nagashima A, Kim Y, Higuchi T, et al. 24-h activity rhythm and sleep in depressed outpatients. *J Psychiatr Res* 2016 Jun;77:27-34. [doi: [10.1016/j.jpsychires.2016.02.022](https://doi.org/10.1016/j.jpsychires.2016.02.022)] [Medline: [26978182](https://pubmed.ncbi.nlm.nih.gov/26978182/)]
22. Tazawa Y, Wada M, Mitsukura Y, Takamiya A, Kitazawa M, Yoshimura M, et al. Actigraphy for evaluation of mood disorders: A systematic review and meta-analysis. *J Affect Disord* 2019 Apr 22;253:257-269. [doi: [10.1016/j.jad.2019.04.087](https://doi.org/10.1016/j.jad.2019.04.087)] [Medline: [31060012](https://pubmed.ncbi.nlm.nih.gov/31060012/)]
23. Tonon AC, Fuchs DFP, Barbosa Gomes W, Levandovski R, Pio de Almeida Fleck M, Hidalgo MPL, et al. Nocturnal motor activity and light exposure: Objective actigraphy-based marks of melancholic and non-melancholic depressive disorder. Brief report. *Psychiatry Res* 2017 Dec;258:587-590. [doi: [10.1016/j.psychres.2017.08.025](https://doi.org/10.1016/j.psychres.2017.08.025)] [Medline: [28844556](https://pubmed.ncbi.nlm.nih.gov/28844556/)]
24. Paudel ML, Taylor BC, Ancoli-Israel S, Stone KL, Tranah G, Redline S, et al. Rest/activity rhythms and cardiovascular disease in older men. *Chronobiol Int* 2011 Apr;28(3):258-266 [FREE Full text] [doi: [10.3109/07420528.2011.553016](https://doi.org/10.3109/07420528.2011.553016)] [Medline: [21452921](https://pubmed.ncbi.nlm.nih.gov/21452921/)]
25. Laposky AD, Bass J, Kohsaka A, Turek FW. Sleep and circadian rhythms: key components in the regulation of energy metabolism. *FEBS Lett* 2008 Jan 09;582(1):142-151 [FREE Full text] [doi: [10.1016/j.febslet.2007.06.079](https://doi.org/10.1016/j.febslet.2007.06.079)] [Medline: [17707819](https://pubmed.ncbi.nlm.nih.gov/17707819/)]
26. Atkinson G, Davenne D. Relationships between sleep, physical activity and human health. *Physiol Behav* 2007 Feb 28;90(2-3):229-235 [FREE Full text] [doi: [10.1016/j.physbeh.2006.09.015](https://doi.org/10.1016/j.physbeh.2006.09.015)] [Medline: [17067643](https://pubmed.ncbi.nlm.nih.gov/17067643/)]
27. Quante M, Mariani S, Weng J, Marinac CR, Kaplan ER, Rueschman M, et al. Zeitgebers and their association with rest-activity patterns. *Chronobiol Int* 2019 Feb;36(2):203-213. [doi: [10.1080/07420528.2018.1527347](https://doi.org/10.1080/07420528.2018.1527347)] [Medline: [30365354](https://pubmed.ncbi.nlm.nih.gov/30365354/)]
28. Finazzi ME, Mesquita ME, Lopes JR, Fu LI, Oliveira MG, Del Porto JA. Motor activity and depression severity in adolescent outpatients. *Neuropsychobiology* 2010;61(1):33-40. [doi: [10.1159/000262178](https://doi.org/10.1159/000262178)] [Medline: [19940518](https://pubmed.ncbi.nlm.nih.gov/19940518/)]
29. Luik AI, Zuurbier LA, Hofman A, van Someren EJW, Tiemeier H. Stability and fragmentation of the activity rhythm across the sleep-wake cycle: the importance of age, lifestyle, and mental health. *Chronobiol Int* 2013 Dec;30(10):1223-1230. [doi: [10.3109/07420528.2013.813528](https://doi.org/10.3109/07420528.2013.813528)] [Medline: [23971909](https://pubmed.ncbi.nlm.nih.gov/23971909/)]

30. Maglione JE, Ancoli-Israel S, Peters KW, Paudel ML, Yaffe K, Ensrud KE, Study of Osteoporotic Fractures Research Group. Depressive symptoms and circadian activity rhythm disturbances in community-dwelling older women. *Am J Geriatr Psychiatry* 2014 Apr;22(4):349-361 [[FREE Full text](#)] [doi: [10.1016/j.jagp.2012.09.003](https://doi.org/10.1016/j.jagp.2012.09.003)] [Medline: [23567424](#)]
31. Difrancesco S, Lamers F, Riese H, Merikangas KR, Beekman ATF, van Hemert AM, et al. Sleep, circadian rhythm, and physical activity patterns in depressive and anxiety disorders: A 2-week ambulatory assessment study. *Depress Anxiety* 2019 Oct;36(10):975-986 [[FREE Full text](#)] [doi: [10.1002/da.22949](https://doi.org/10.1002/da.22949)] [Medline: [31348850](#)]
32. Kim H, Lee S, Lee S, Hong S, Kang H, Kim N. Depression prediction by using Ecological Momentary Assessment, actiwatch data, and machine learning: observational study on older adults living alone. *JMIR Mhealth Uhealth* 2019 Oct 16;7(10):e14149 [[FREE Full text](#)] [doi: [10.2196/14149](https://doi.org/10.2196/14149)] [Medline: [31621642](#)]
33. Penninx BWJH, Beekman ATF, Smit JH, Zitman FG, Nolen WA, Spinhoven P, NESDA Research Consortium. The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int J Methods Psychiatr Res* 2008;17(3):121-140 [[FREE Full text](#)] [doi: [10.1002/mpr.256](https://doi.org/10.1002/mpr.256)] [Medline: [18763692](#)]
34. Booij SH, Bos EH, Bouwmans MEJ, van Faassen M, Kema IP, Oldehinkel AJ, et al. Cortisol and α -amylase secretion patterns between and within depressed and non-depressed individuals. *PLoS One* 2015 Jul 6;10(7):e0131002 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0131002](https://doi.org/10.1371/journal.pone.0131002)] [Medline: [26148294](#)]
35. Schoevers RA, van Borkulo CD, Lamers F, Servaas M, Bastiaansen JA, Beekman ATF, et al. Affect fluctuations examined with ecological momentary assessment in patients with current or remitted depression and anxiety disorders. *Psychol Med* 2020 Apr 01:1-10. [doi: [10.1017/S0033291720000689](https://doi.org/10.1017/S0033291720000689)] [Medline: [32234092](#)]
36. Wittchen HU. Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): a critical review. *J Psychiatr Res* 1994;28(1):57-84. [doi: [10.1016/0022-3956\(94\)90036-1](https://doi.org/10.1016/0022-3956(94)90036-1)] [Medline: [8064641](#)]
37. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med* 1996 May;26(3):477-486. [doi: [10.1017/s0033291700035558](https://doi.org/10.1017/s0033291700035558)] [Medline: [8733206](#)]
38. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio, TX: Psychological Corporation; 1996:1-82.
39. Eslinger DW, Rowlands AV, Hurst TL, Catt M, Murray P, Eston RG. Validation of the GENEActiv Accelerometer. *Med Sci Sports Exerc* 2011 Jun;43(6):1085-1093. [doi: [10.1249/MSS.0b013e31820513be](https://doi.org/10.1249/MSS.0b013e31820513be)] [Medline: [21088628](#)]
40. Pavey TG, Gomersall SR, Clark BK, Brown WJ. The validity of the GENEActiv wrist-worn accelerometer for measuring adult sedentary time in free living. *J Sci Med Sport* 2016 May;19(5):395-399. [doi: [10.1016/j.jsams.2015.04.007](https://doi.org/10.1016/j.jsams.2015.04.007)] [Medline: [25956687](#)]
41. Hager ER, Treuth MS, Gormely C, Epps L, Snitker S, Black MM. Ankle accelerometry for assessing physical activity among adolescent girls: threshold determination, validity, reliability, and feasibility. *Res Q Exerc Sport* 2015;86(4):397-405 [[FREE Full text](#)] [doi: [10.1080/02701367.2015.1063574](https://doi.org/10.1080/02701367.2015.1063574)] [Medline: [26288333](#)]
42. Stavrakakis N, Booij SH, Roest AM, de Jonge P, Oldehinkel AJ, Bos EH. Temporal dynamics of physical activity and affect in depressed and nondepressed individuals. *Health Psychol* 2015 Dec;34S:1268-1277. [doi: [10.1037/hea0000303](https://doi.org/10.1037/hea0000303)] [Medline: [26651468](#)]
43. Bouwmans MEJ, Bos EH, Booij SH, van Faassen M, Oldehinkel AJ, de Jonge P. Intra- and inter-individual variability of longitudinal daytime melatonin secretion patterns in depressed and non-depressed individuals. *Chronobiol Int* 2015 Apr 27;32(3):441-446. [doi: [10.3109/07420528.2014.973114](https://doi.org/10.3109/07420528.2014.973114)] [Medline: [25347155](#)]
44. American Psychiatric Association: Diagnostic and statistical manual of mental disorders, fourth edition, text revision. Washington, DC: American Psychiatric Association; 2000.
45. Vallance JK, Winkler EA, Gardiner PA, Healy GN, Lynch BM, Owen N. Associations of objectively-assessed physical activity and sedentary time with depression: NHANES (2005-2006). *Prev Med* 2011 Oct;53(4-5):284-288. [doi: [10.1016/j.ypmed.2011.07.013](https://doi.org/10.1016/j.ypmed.2011.07.013)] [Medline: [21820466](#)]
46. Robillard R, Naismith SL, Smith KL, Rogers NL, White D, Terpening Z, et al. Sleep-wake cycle in young and older persons with a lifetime history of mood disorders. *PLoS One* 2014;9(2):e87763 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0087763](https://doi.org/10.1371/journal.pone.0087763)] [Medline: [24586290](#)]
47. Sadeh A, Alster J, Urbach D, Lavie P. Actigraphically based automatic bedtime sleep-wake scoring: Validity and clinical applications. *J Ambul Monit* 1989;2(3):209-216.
48. Marler MR, Gehrman P, Martin JL, Ancoli-Israel S. The sigmoidally transformed cosine curve: a mathematical model for circadian rhythms with symmetric non-sinusoidal shapes. *Stat Med* 2006 Nov 30;25(22):3893-3904. [doi: [10.1002/sim.2466](https://doi.org/10.1002/sim.2466)] [Medline: [16381069](#)]
49. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 2012 May 18;36(1):27-46. [doi: [10.1111/j.1600-0587.2012.07348.x](https://doi.org/10.1111/j.1600-0587.2012.07348.x)]
50. Field A. *Discovering Statistics Using IBM SPSS Statistics*. Los Angeles, CA: Sage Publications Ltd; 2013.
51. Piccinelli M, Wilkinson G. Gender differences in depression. Critical review. *Br J Psychiatry* 2000 Dec;177:486-492. [doi: [10.1192/bjp.177.6.486](https://doi.org/10.1192/bjp.177.6.486)] [Medline: [11102321](#)]
52. Miech RA, Shanahan MJ. Socioeconomic status and depression over the life course. *J Health Soc Behav* 2000 Jun;41(2):162-176. [doi: [10.2307/2676303](https://doi.org/10.2307/2676303)]

53. Steyerberg EW, Eijkemans MJ, Harrell FE, Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. *Stat Med* 2000 Apr 30;19(8):1059-1079. [doi: [10.1002/\(sici\)1097-0258\(20000430\)19:8<1059::aid-sim412>3.0.co;2-0](https://doi.org/10.1002/(sici)1097-0258(20000430)19:8<1059::aid-sim412>3.0.co;2-0)] [Medline: [10790680](#)]
54. Harrell FE. Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis. Springer International Publishing AG Switzerland: Springer Ser Stat; 2015.
55. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010 Jan;21(1):128-138 [FREE Full text] [doi: [10.1097/EDE.0b013e3181c30fb2](https://doi.org/10.1097/EDE.0b013e3181c30fb2)] [Medline: [20010215](#)]
56. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015 Jan 06;162(1):W1-73 [FREE Full text] [doi: [10.7326/M14-0698](https://doi.org/10.7326/M14-0698)] [Medline: [25560730](#)]
57. Kim J, Nakamura T, Yamamoto Y. A momentary biomarker for depressive mood. *In Silico Pharmacol* 2016 Dec;4(1):4 [FREE Full text] [doi: [10.1186/s40203-016-0017-6](https://doi.org/10.1186/s40203-016-0017-6)] [Medline: [26979449](#)]
58. de Jonge M, Dekker JJM, Kikkert MJ, Peen J, van Rijsbergen GD, Bockting CLH. The role of affect in predicting depressive symptomatology in remitted recurrently depressed patients. *J Affect Disord* 2017 Mar 01;210:66-71. [doi: [10.1016/j.jad.2016.12.015](https://doi.org/10.1016/j.jad.2016.12.015)] [Medline: [28013124](#)]
59. Wichers M, Peeters F, Geschwind N, Jacobs N, Simons CJP, Derom C, et al. Unveiling patterns of affective responses in daily life may improve outcome prediction in depression: a momentary assessment study. *J Affect Disord* 2010 Jul;124(1-2):191-195. [doi: [10.1016/j.jad.2009.11.010](https://doi.org/10.1016/j.jad.2009.11.010)] [Medline: [20004977](#)]
60. Wichers M, Lothmann C, Simons CJP, Nicolson NA, Peeters F. The dynamic interplay between negative and positive emotions in daily life predicts response to treatment in depression: a momentary assessment study. *Br J Clin Psychol* 2012 Jun;51(2):206-222. [doi: [10.1111/j.2044-8260.2011.02021.x](https://doi.org/10.1111/j.2044-8260.2011.02021.x)] [Medline: [22574805](#)]
61. Geschwind N, Nicolson NA, Peeters F, van Os J, Barge-Schaapveld D, Wichers M. Early improvement in positive rather than negative emotion predicts remission from depression after pharmacotherapy. *Eur Neuropsychopharmacol* 2011 Mar;21(3):241-247 [FREE Full text] [doi: [10.1016/j.euroneuro.2010.11.004](https://doi.org/10.1016/j.euroneuro.2010.11.004)] [Medline: [21146375](#)]
62. Cohen JR, Thakur H, Young JF, Hankin BL. The development and validation of an algorithm to predict future depression onset in unselected youth. *Psychol Med* 2019 Oct 02;1-9. [doi: [10.1017/S0033291719002691](https://doi.org/10.1017/S0033291719002691)] [Medline: [31576786](#)]
63. Andrews JA, Harrison RF, Brown LJE, MacLean LM, Hwang F, Smith T, et al. Using the NANA toolkit at home to predict older adults' future depression. *J Affect Disord* 2017 Apr 15;213:187-190 [FREE Full text] [doi: [10.1016/j.jad.2017.02.019](https://doi.org/10.1016/j.jad.2017.02.019)] [Medline: [28259086](#)]
64. Pyszczynski T, Hamilton JC, Herring FH, Greenberg J. Depression, self-focused attention, and the negative memory bias. *J Pers Soc Psychol* 1989 Mar;57(2):351-357 [FREE Full text] [doi: [10.1037/0022-3514.57.2.351](https://doi.org/10.1037/0022-3514.57.2.351)]
65. Karas M, Bai J, Strączkiewicz M, Harezlak J, Glynn NW, Harris T, et al. Accelerometry data in health research: challenges and opportunities. *Stat Biosci* 2019 Jul;11(2):210-237 [FREE Full text] [doi: [10.1007/s12561-018-9227-2](https://doi.org/10.1007/s12561-018-9227-2)] [Medline: [31762829](#)]
66. Reichert M, Lutz A, Deuschle M, Gilles M, Hill H, Limberger MF, et al. Improving motor activity assessment in depression: which sensor placement, analytic strategy and diurnal time frame are most powerful in distinguishing patients from controls and monitoring treatment effects. *PLoS One* 2015;10(4):e0124231 [FREE Full text] [doi: [10.1371/journal.pone.0124231](https://doi.org/10.1371/journal.pone.0124231)] [Medline: [25885258](#)]
67. Todder D, Caliskan S, Baune BT. Longitudinal changes of day-time and night-time gross motor activity in clinical responders and non-responders of major depression. *World J Biol Psychiatry* 2009;10(4):276-284. [doi: [10.3109/15622970701403081](https://doi.org/10.3109/15622970701403081)] [Medline: [19921969](#)]
68. Smagula SF, Boudreau RM, Stone K, Reynolds CF, Bromberger JT, Ancoli-Israel S, Osteoporotic Fractures in Men (MrOS) Research Group. Latent activity rhythm disturbance sub-groups and longitudinal change in depression symptoms among older men. *Chronobiol Int* 2015;32(10):1427-1437 [FREE Full text] [doi: [10.3109/07420528.2015.1102925](https://doi.org/10.3109/07420528.2015.1102925)] [Medline: [26594893](#)]
69. McClung CA. Circadian rhythms in mood disorders. In: *Circadian Medicine*. Hoboken, NJ: John Wiley & Sons, Inc; 2015.
70. McClung CA. How might circadian rhythms control mood? Let me count the ways. *Biol Psychiatry* 2013 Aug 15;74(4):242-249 [FREE Full text] [doi: [10.1016/j.biopsych.2013.02.019](https://doi.org/10.1016/j.biopsych.2013.02.019)] [Medline: [23558300](#)]
71. Lyall LM, Wyse CA, Graham N, Ferguson A, Lyall DM, Cullen B, et al. Association of disrupted circadian rhythmicity with mood disorders, subjective wellbeing, and cognitive function: a cross-sectional study of 91 105 participants from the UK Biobank. *Lancet Psychiatry* 2018 Jun;5(6):507-514. [doi: [10.1016/S2215-0366\(18\)30139-1](https://doi.org/10.1016/S2215-0366(18)30139-1)] [Medline: [29776774](#)]
72. Nebes RD, Buysse DJ, Halligan EM, Houck PR, Monk TH. Self-reported sleep quality predicts poor cognitive performance in healthy older adults. *J Gerontol B Psychol Sci Soc Sci* 2009 Mar;64(2):180-187 [FREE Full text] [doi: [10.1093/geronb/gbn037](https://doi.org/10.1093/geronb/gbn037)] [Medline: [19204069](#)]
73. Toll DB, Janssen KJM, Vergouwe Y, Moons KGM. Validation, updating and impact of clinical prediction rules: a review. *J Clin Epidemiol* 2008 Nov;61(11):1085-1094. [doi: [10.1016/j.jclinepi.2008.04.008](https://doi.org/10.1016/j.jclinepi.2008.04.008)] [Medline: [19208371](#)]

74. Mata J, Thompson RJ, Jaeggi SM, Buschkuhl M, Jonides J, Gotlib IH. Walk on the bright side: physical activity and affect in major depressive disorder. *J Abnorm Psychol* 2012 May;121(2):297-308 [FREE Full text] [doi: [10.1037/a0023533](https://doi.org/10.1037/a0023533)] [Medline: [21553939](https://pubmed.ncbi.nlm.nih.gov/21553939/)]
75. Thompson RJ, Mata J, Jaeggi SM, Buschkuhl M, Jonides J, Gotlib IH. The everyday emotional experience of adults with major depressive disorder: Examining emotional instability, inertia, and reactivity. *J Abnorm Psychol* 2012 Nov;121(4):819-829 [FREE Full text] [doi: [10.1037/a0027978](https://doi.org/10.1037/a0027978)] [Medline: [22708886](https://pubmed.ncbi.nlm.nih.gov/22708886/)]
76. Bylsma LM, Taylor-Clift A, Rottenberg J. Emotional reactivity to daily events in major and minor depression. *J Abnorm Psychol* 2011 Feb;120(1):155-167. [doi: [10.1037/a0021662](https://doi.org/10.1037/a0021662)] [Medline: [21319928](https://pubmed.ncbi.nlm.nih.gov/21319928/)]
77. Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle NC, Mora-Gonzalez J, Löf M, et al. Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports Med* 2017 Sep;47(9):1821-1845. [doi: [10.1007/s40279-017-0716-0](https://doi.org/10.1007/s40279-017-0716-0)] [Medline: [28303543](https://pubmed.ncbi.nlm.nih.gov/28303543/)]
78. van Os J, Verhagen S, Marsman A, Peeters F, Bak M, Marcelis M, ESM-MERGE Investigators PhD, et al. The experience sampling method as an mHealth tool to support self-monitoring, self-insight, and personalized health care in clinical practice. *Depress Anxiety* 2017 Dec;34(6):481-493. [doi: [10.1002/da.22647](https://doi.org/10.1002/da.22647)] [Medline: [28544391](https://pubmed.ncbi.nlm.nih.gov/28544391/)]
79. Bos FM, Snippe E, Bruggeman R, Wichers M, van der Krieke L. Insights of patients and clinicians on the promise of the experience sampling method for psychiatric care. *Psychiatr Serv* 2019 Nov 01;70(11):983-991. [doi: [10.1176/appi.ps.201900050](https://doi.org/10.1176/appi.ps.201900050)] [Medline: [31434558](https://pubmed.ncbi.nlm.nih.gov/31434558/)]

Abbreviations

AC: activity counts

AIC: Akaike information criterion

AUC: the area under the receiver operating characteristic curve

BDI-II: Beck Depression Inventory-II

CIDI: Composite International Diagnostic Interview

DSM: Diagnostic and Statistical Manual of Mental Disorders

ENMO: Euclidian Norm Minus One ESM: experience sampling method

IDS-SR: Inventory of Depressive Symptomatology, self-report

MESOR: midline estimating statistic of rhythm

MOOVD: Mood and Movement in Daily Life

MVPA: moderate-to-vigorous physical activity

NESDA: the Netherlands Study of Depression and Anxiety

NESDA-EMAA: Ecological Momentary Assessment and Actigraphy substudy of NESDA

RAR: rest-activity rhythm

TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis

TST: total sleep time

VIF: variance inflation factor

Edited by G Eysenbach; submitted 19.07.20; peer-reviewed by J Li, Y Li; comments to author 28.07.20; revised version received 13.08.20; accepted 26.10.20; published 01.12.20

Please cite as:

Minaeva O, Riese H, Lamers F, Antypa N, Wichers M, Booij SH

Screening for Depression in Daily Life: Development and External Validation of a Prediction Model Based on Actigraphy and Experience Sampling Method

J Med Internet Res 2020;22(12):e22634

URL: <https://www.jmir.org/2020/12/e22634>

doi: [10.2196/22634](https://doi.org/10.2196/22634)

PMID:

©Olga Minaeva, Harriëtte Riese, Femke Lamers, Niki Antypa, Marieke Wichers, Sanne H Booij. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 01.12.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.